DOCUMENT RESUME

ED 037 502 UD 009 698

AUTHOR Gordon, Edmund W.
TITLE Problems in the Assessment of Intermediate-Range Effects of Head Start Programs.
PUB DATE May 68
NOTE 22p.; Paper presented at Educational Testing Service Meeting

EDRS PRICE EDRS Price MF-$0.25 HC-$1.20
DESCRIPTORS Affective Behavior, *Cognitive Development, Compensatory Education, *Disadvantaged Youth, Evaluation Criteria, Evaluation Methods, Intervention, Preschool Children, *Preschool Evaluation, *Preschool Programs, Research Methodology, *Research Problems, Student Motivation, Teacher Influence
IDENTIFIERS Head Start Programs

ABSTRACT

More carefully controlled research distinguishing more facets which may influence cognitive and social performance in young children is considered necessary before Head Start or any preschool program can be adequately assessed. Shortcomings in evaluation methodology discussed are as follows: evaluation research has not addressed itself to program input variables; preschool children are often evaluated as a homogeneous population; a disproportionate emphasis has been placed on cognitive development while ignoring the equally important and interrelated affective domain; studies of cognitive development have been too dependent on intelligence and achievement tests; and, preschool programs have been relied on too heavily as the only means of intervention. Wolf and Stein's (1966) follow-up research project on the effects of Head Start programs for Puerto-Rican, black, and non-Puerto-Rican white children in New York City is presented as a study which raises important issues in research methodology. Reviews of studies of cognitive development are presented, which support the theory that the degree of progress made by Head Start graduates and children without preschool experience can invariably be related to a child's base level of function. Problems of evaluating teacher influence and other variables of the affective domain are also discussed. (KG)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED037502

Problems in the Assessment of Intermediate-Range Effects of Head Start Programs*

Edmund W. Gordon, Ed. D.
Ferkauf Graduate School of Humanities and Social Sciences, Yeshiva University

May 1968

Long before Head Start, the relationship between institutionalized efforts to modify the child socially and intellectually and his subsequent performance in these areas had been an important research issue in educational psychology. A number of studies had tried to document both the immediate and long-range effects of preschool education programs, particularly in relation to the development of middle-class children. More recently, studies have become concerned with the development of lower-class children. While many investigations have noted dramatic changes immediately after the preschool program (such as a 10-point increase in intelligence test scores after a few weeks), follow-up studies present a less consistent view: some of them point to significant leads by preschool graduates over their peers as far removed from the experience as the third grade, while others report the preschoolers to be indistinguishable from their classmates in intellectual or social achievement after only a few months in public school.

UD 009 698

*Paper presented at Educational Testing Service meeting.

Although Head Start was initiated on the basis of perceived social need, logical expectations and some theoretical rationale, there was and still is no conclusive evidence to substantiate the measurable benefits of preschool intervention for either middle-class or lower-class children. However, with the inauguration of the Project, investigators have at last been given a massive preschool program in which to conduct such research in relation to the disadvantaged child. Research with this additional population has inherent in it many of the difficulties which were unresolved in earlier investigations on middle-class children.

Problems in Methodology

A number of factors are responsible for the inconclusive results which have until now characterized research in this area. First, preschool models have tended to be lumped together rather than evaluated according to their specific content. While vast differences exist in program philosophy, quality, length and content, most evaluative research has paid little attention to these input variables. Second, despite well-documented variations in functional characteristics within a single socio-economic or ethnic group, research has usually been conducted as if a homogeneous population of preschoolers were being viewed. Third, most studies have focused on cognitive development and have ignored the equally important and interrelated affective domain. Fourth, studies of cognitive development have been dependent upon intelligence and achievement tests, some of which may measure too limited a range and quality of behaviors to reflect accurately the status of the children

studied. Finally, while preschool intervention may facilitate a child's adjustment to, and progress in, public school, and may come to be regarded as a necessity for some children, it cannot by itself ensure this development. A large measure of the developmental course followed depends on the prior conditions and subsequent environmental influences confronting the child.

The Meaning of Change in Tested Functioning

One problem which has received considerable attention in evaluation research studies of the preschool experience even before Head Start has been the meaning of the changes observed in participating children. A number of interpretations have been given to increases in test scores which occur between the beginning and end of intervention programs. Those who are most enthusiactic readily claim that the preschool program has simply improved intellectual capacity, or at least intellectual function in the children. The more skeptical explain most of the test-retest change in scores on the basis of the practice effect--greater familiarity with the test items and test situation. Weikart (1964) has pointed to the familiarity with the test materials which children gain during the preschool experience. His position is close to those who could explain the changes in pre-post test scores on the basis of accidental if not intentional coaching in test taking. The child's greater ease with middle-class teachers and test administrators as a result of the preschool experience has also been noted by several observers. In another context, Zigler (1964) has suggested that pre-post intelligence test score changes may more greatly

reflect improved motivation and task involvement than changes in intellectual functioning.

Variables Contributing to Permanence of Gains

Although the reason for improvement in test scores immediately following an intervention program may be arguable, the fact of improvement is almost universally observed. It is the maintenance of these gains which is far less frequently reported. The consolidation of qualitative and quantitative changes in intellectual function, and their transfer to other situations is, in fact, one of the most acute challenges which educators face. In discussions of Head Start research, there has been much concern over a "levelling off" of the children's gains, which appears at times to occur after several months in public school. Some critics have mislabeled this phenomenon a "fade-out" effect, implying that the preschool experience becomes inconsequential or is lost after a time, and that the child might have been as well off without it. However, it is known that periods of rapid achievement are usually followed by periods of consolidation during which gains are negligible. Such a levelling off has been observed in a variety of learning situations with both middle-class and lower-class children, and it is often followed at some later date by further gains. The critical issue is not only the constancy of the gains or advantage that the Head Start child may have; rather an equally important issue involves the by-products of even a temporary advantage which preschool children can have over youngsters who are not given such an experience.

It must be noted that research on preschool intervention for the disadvantaged is only beginning to approach a level of sophistication which reflects the intricacies of social and conceptual learning in children. In the rush to make data available on Head Start programs, little attention has been given to the inter-related aspects of change which occur during the preschool period. The relative newness of the program has also limited the length of follow-up studies. Of those reported, progress has been measured no later than 6 to 8 months after intervention. Moreover, the largest number of studies available on Head Start still come from the first eight-week program held in the summer of 1965. Despite many weaknesses in the research, preliminary indications are not discouraging. If the results are at times contradictory, they point to the richness of the subject matter for research, and underscore the fact that no program as extensive and varied as this one should be judged purely on the basis of these initial efforts at evaluation.

An important study which raises many of the issues involved in gauging the effect of Head Start is the research project conducted by Wolff and Stein (1966) on Puerto Rican, Negro and non-Puerto Rican white children in New York City. These children had attended Head Start programs for at least six weeks in summer, 1965. Popularized accounts of this study emphasized the decreasing advantage in functional levels which the authors found Head Start children to have over their peers as the kindergarten year progressed. After an initial period in which the Head Start children had stood out because of their more rapid adaptation, most kindergarten teachers felt that they could no longer observe any superiority in

these children. At the end of six months, Head Start graduates also showed no advantages in Preschool Inventory scores over their classmates who had had no pre-school experience. However, the widely reported "levelling effect" was only one aspect of the Wolff-Stein investigation. The authors" discussions of the complexities of this levelling effect, as well as several other interesting findings, have largely been ignored. Despite some methodological flaws, the study demands further attention.

The Wolff-Stein sample consisted of 567 kindergarten children: 179 were Head Start graduates, 388 had never attended a preschool program. The two groups were theoretically matched for socio-economic and ethnic backgrounds, level of intelligence, age and sex. A major weakness of the study is the inequality of the socio-economic status of the comparison groups used. Only half as many of the control group as opposed to the experimental Head Start group came from families in dire poverty ($3000 or less annually) and fewer were on welfare. Non-Head Start families also tended to be smaller, non-Head Start parents had more education than parents of Head Start children, although both groups professed to have similar educational goals for their children. The matched children did live in the same neighborhoods. Both groups attended the morning and afternoon kindergarten sessions of four public elementary schools: one Negro, two Puerto Rican and one which served children of ethnically and socially mixed backgrounds.

In contrast to the majority of Head Start follow-up investigations which concentrated on standardized tests to ascertain achievement levels, Wolff and Stein used

assessments by the fifteen teachers who were in charge of the 30 kindergarten classrooms. Teachers were given ratings and questionnaires to fill out, and were observed as they functioned in their classes. Evaluations of the children included such items as their initial and long-term (6 months) adjustment to school, their behavior towards other children and the teacher, their work and listening habits, and their mastery of concepts necessary for first grade work.

An important finding in this study is that ratings of both Head Start and non-Head Start children were highly correlated with the ethnic character of the schools they attended.* In the Negro and both Puerto Rican schools, the teachers ranked more Head Start graduates than non-Head Start children in the top 30 per cent of the class in readiness to enter first grade. Only in the mixed neighborhood school, which contained 28.3% white, middle-class children in contrast to between zero and 6% in the other schools, did the teachers give lower rankings to Head Start graduates in this category. This exception parallels a difference in emphasis which the authors found among the teachers in the four schools: while teachers in the mixed neighborhood school gave more weight in their ratings to "learning concepts," teachers in the Negro and Puerto Rican schools concentrated on "social factors" or "readiness." In all three minority-group schools, observers found much emphasis on getting ready for "the school experience," but little on actual instruction and learning. The discrepancy in emphasis supports a common observation that one cause

---

* The lack of specific information on the content of the preschool programs themselves may, of course, have contributed to this outcome. As in the following studies mentioned in this report, the fact of having had a "Head Start experience" is assumed as a uniform variable with no further differentiation.

of low achievement among children in the city ghetto schools is the low expectations of their teachers (Gordon & Wilkerson, 1966).

On a blind evaluation of all children in their classes, the 15 teachers pointed to Head Start children as having made a more rapid social adjustment than children who entered without preschool experience, although the teachers' evaluations of long-term adjustment once again followed school lines. Teachers in the school which served the ethnically and socio-economically mixed group were less enthusiastic about the long-range behavioral advantages resulting from Head Start than were those in the Negro and Puerto Rican schools. While teachers in the mixed neighborhood school could see no difference between Head Start children and children without preschool experience in their ability to relate to peers, teachers in the Negro and Puerto Rican schools thought Head Start graduates related better to others than did their classmates. The largest advantaged given to the Head Start children over their non-Head Start peers occurred in the two Puerto Rican schools.

Ratings of the children's behavior toward the teacher herself were not linked to the ethnic character of the school, but appeared to vary with individual teachers. None out of 12 teachers who responded to a question on the effect of Head Start in this area felt that the experience had helped the children to relate to them. One of the best teachers (as judged by observers) remarked perceptively that, "If the child had had a poor Head Start teacher he would have a harder time relating to a kindergarten teacher than a child with no previous school at all." That this relationship exists is evidenced by Wolff's and Stein's additional finding that Head Start children were more affected than their lower-class peers by the

quality of their kindergarten teacher. Those Head Start graduates with teachers whom observers had rated "good" scored higher than their non-Head Start peers on the Pre-school Inventory, while those with kindergarten teachers whom observers had ranked "poor" did more poorly than the children with no preschool experience. In the light of Coleman's finding (1966) that disadvantaged children are more affected by quality of teaching than are advantaged children, we can only assume that sensitivity to teacher quality is even further heightened by early school experience. It may therefore become considerably more crucial to raise the quality of teachers in disadvantaged neighborhoods, as children's sensitivity to their school experience is heightened by programs such as Head Start.

Several other teacher and classroom variables must be considered when gauging the effectiveness of any preschool program in helping its graduates to succeed in the primary grades. A particularly provocative finding in the Wolff-Stein report is that the average or better-than average kindergarten teacher concentrated two-thirds of her attention on the top half of the ranked class, whether the interaction was praise, scolding, or instruction. One the other hand, teachers whom observers rated poor gave most of their attention to the bottom half, with 50-75% in the form of negative comments. This finding raises particular concern for the lower-class minority group youngsters. Teachers in schools serving disadvantaged populations are generally rated inferior both in educational background and orientation to those in schools serving integrated or white middle-class populations (Coleman, 1966). Thus the plight of the pupil in the bottom-half of the class is likely to be made even more difficult in these schools.

While the focus of teacher attention on either the top or bottom section of the class might appear to argue for more homogeneous classrooms, there is little to guarantee that a teacher will not make her own divisions even with an initially homogeneous group of children. This is particularly true when the classroom is grossly overcrowded, as are the majority in the schools serving disadvantaged areas. Additional support for heterogeneous grouping, however, comes from the finding that achievement increases for lower-class students as the ratio of middle-class to lower-class children goes up. (U.S. Commission on Civil Rights, 1967.) An economically homogeneous class thus enhances the chance that the disadvantaged child will be faced with multiple school-based handicaps.

One of the major problems in any minority-group school may relate to background differences between the teachers and their pupils. These differences are usually ethnic and socio-economic, and often result in gulfs in style and temperament between the teacher and the class. In the Wolff and Stein sample, with the exception of one Negro teacher, all were non-Puerto Rican whites. Despite the fact that over 330 of the kindergarten children were Puerto Rican, only one teacher of the 15 spoke Spanish. In nine of the 15 classrooms there were no books, pictures or other materials which related to the minority group children's background. Finally, four of the teachers were observed to be prejudiced or to hold stereotyped attitudes towards their children, and even those teachers rated "good" tended to favor pupils whose styles of work resembled their own. These findings may help to explain the advantage which non-Head Start, middle-class children held over their Head Start peers in the mixed-neighborhood school.

In addition to the factors related to the quality and attitudes of the kindergarten teacher, Wolff and Stein found a crucial element in the retention of advantages derived from preschool to be the proportion of Head Start children in the kindergarten. Classroom composition appeared to affect both the students' progress through the curriculum and the teachers' evaluations of the effectiveness of Head Start. All teachers with classrooms composed of 50 per cent or more Head Start children thought that the summer program had helped the individual child's adjustment, and noted that they had been able to move more rapidly through the curriculum than in previous years with children of similar backgrounds. When the classes were comprised of less than 25 per cent Head Start graduates, however, the teachers failed to see any difference which the program might have made either for the individual child or for the class as a whole. Thus, having a high percentage of Head Start graduates in the classroom functioned to raise the level of achievement in much the same way as having a larger percentage of middle-class children. These findings are doubly interesting, since they have implications for influencing developmental rate and also serve to confound long-term follow-up evaluations where controls are drawn from the same classrooms.

Several additional sources of confusion to follow-up findings have been noted in the literature on preschooling for the disadvantaged. Susan Gray (1966) reports the confounding impact of preschool on the children's families and even on the community in which the children lived. Reporting on the Early Training Project in Nashville, Tennessee, in the summers between 1963 and 1965, Grey states that at the end of each school year the controls in the same class had caught up to the

gains made during the summer by her experimental group. However, another control group in a town 60 miles away did not show the same progress. In addition, the younger brothers and sisters of the children in a preschool program have been observed by Kohlberg (1966) to be making unusual progress probably as a result of the influence of the program on their parents. When they were studied, their siblings were found to be on a higher level than their older siblings had been, as well as better prepared than their peers who had had no older siblings in the program.

This diffusion of gains brings into question the validity of any study which attempts to determine long-range benefits of a special training program solely on the basis of comparison between the experimental group and a control group with whom it has regular contact. Unfortunately, most investigators of Head Start children have assumed that in order to ascertain the effects of compensatory education, the experimental group would have to be compared to a control group from a comparable socio-economic background who had had as nearly as possible the same educational experiences--with the exception of the one training period under study. Children in the same classroom were most often considered the ideal solution. That in many cases differences between Head Start children and their matched classroom peers are still found at the end of six to eight months of contact is a testimony to the strength of some of these preschool programs.

Problems in Evaluating Cognitive Gains

Despite the alleged broader focus of Head Start, the strikingly low level of conceptual facility which lower-class children show in comparison to children from

the middle-class has made the development of cognitive skills an important variable in Head Start programs and has come to dominate evaluation of them. In nearly all studies contrasting intelligence and achievement ratings before and immediately after the Head Start experience, significant improvements are found in the experimental group which go far beyond developmental changes due to increased age alone. In contrast to these striking short-term gains, the follow-up studies which consider the longitudinal development of Head Start populations reflect more ambiguous change. In most studies, the experimental group children show less consistent gain or maintence of advantage than do control groups with no preschool experience.

Three generalizations can be drawn from the results of the dozen or so follow-up investigations on intelligence and achievement levels which have been completed to date. First, the control group of lower-income kindergarten children without preschool experience usually makes more rapid progress from their lower starting point than the Head Start group.* Second, while the Head Start group's progress during the first year in school is not as rapid as that of the control group, it most often manages to stay ahead in the cognitive variables measured. When developmental level is controlled, as in a study by Spigel and McBane (1966), test scores show Head Start children making greater cognitive gains during the school year. Third, both Head Start graduates and their matched non-Head Start peers remain significantly behind the middle-class children in the developmental skills being measured.

*This may also be said for first-grade children in regions where Head Start has functioned as a substitute for kindergarten.
functioned as a substitute for kindergarten.

Thus intelligence and achievement tests administered toward the end of the first public school year show middle-class children ahead, followed after a significant gap by lower-class Head Start graduates, and finally by lower-class children with no preschool experience.

In both Head Start graudates and children without preschool, the degree of progress made during kindergarten can be inversely related to the child's base level of function. In other words, with the exception of mentally defective subjects, those children who begin school with the lowest level of function usually make the greatest amount of progress during the school year. However, in most cases their ultimate achievement is still significantly behind those with higher base levels. This differential rate of response to the kindergarten experience has been demonstrated with respect to socio-economic class, ethnic group and sex, as well as initial intelligence test level. It may be related to at least four factors: (1) some, but certainly not all, of this change may be accounted for on the basis of extreme scorers to gravitate toward the mean on retest; (2) the tests may be designed to tap changes in lower levels of function more readily than in higher levels of function; (3) children who test lower initially may indeed have a greater room for change, although their optimum functioning level is still comparatively low; and (4) the Head Start and/or kindergarten experience may often speak more appropriately to the needs of children with lower levels of function.

A study by Chesteen (1966) shows differential rate of improvement following the Head Start experience to be related to race and socio-economic class. Chesteen

compared 81 Head Start children with 126 lower-class, classroom peers and 28 middle-class children in Baton Rouge, Louisiana. All groups contained both Negro and white children; only the Head Start graduates had had any preschool experience. Since the Head Start centers in Baton Rouge had accepted some middle-class children, the Head Start sample was also divided according to socio-economic groups. The children were tested in September (immediately following the preschool program), at mid-term, and at the end of their kindergarten year, on the Stanford-Binet IQ test and the Primary Mental Abilities Test of school readiness and achievement.

The Head Start children sustained their gains of the previous summer, even increasing their scores slightly; at the end of the year they were ahead of the disadvantaged control group--but still significantly behind the middle-class children. Of the Head Start graduates, low-income white children had made greater gains during the summer than low-income Negro children, but during the kindergarten year the low-income Negro children made more progress, apparently manifesting their advances of the previous summer but manifesting them late. While these lower-class Negro children never caught up to the other groups, they increased their PMA scores three times as much as did the middle-class children, who probably began the year working closer to their optimal level.

Sigel and McBane's investigation of cognitive competence and symbolization levels among five-year olds focused its attention on lower and middle-class Negro children in Detroit. Differences were revealed in cognitive growth between socio-economic groups as well as between the sexes within each group. A sample of

children from the summer 1965 Head Start programs was compared with two groups who had not had preschool training: viz., lower-class children from the same neighborhoods, and middle-class children. The authors used a series of tests--including a categorization test of style and consistency, a block design test, a motor inhibition test, and the Peabody Picture Vocabulary Test--to determine the developmental levels of these children. Testing was conducted 2-3 months after the beginning of kindergarten and again 6-8 months later.

In style and consistency of categorization, both Head Start children and their lower-class controls began the year at approximately the same level--significantly below the middle-class group, however. During the school year, however, the Head Start children made greater progress than their lower-class controls, and at the end testing they approached the scores of the middle-class controls. On the PPVT, where the Head Starters had initially tested much higher than their lower-class peers, the gap between the two groups narrowed during the year, but the middle-class group remained significantly ahead of both lower-class groups.

Middle-class and lower-class Head Start girls were developmentally behind the boys of their respective socio-economic groups at the fall testing on the PPVT and other tests of conceptual function. During the year, they made greater progress than the boys in their socio-economic groups, and on several of the tests were at least on the same level by the time of final examination.

Several studies point to the conclusion that without regard to ethnic, sexual or socio-economic group, within certain limits, the greater the initial degree of intellectual subnormality (as opposed to mental defect), the more likely is there to be significant gain during the year following Head Start. An investigation by

Krider and Petsch (1966) of Head Start graduates in Lincoln, Nebraska, showed that the base-line intelligence level of the children correlated with the amount of progress which could be shown to have occurred eight months after the Head Start experience. When the authors compared scores of Head Start children with those of matched controls on the Stanford-Binet and the Preschool Inventory in September and spring of their kindergarten year, there appeared to be no significant difference in the progress which the two groups had made. However, when the two groups of children were subdivided according to intelligence levels, it was found that those Head Start children classified as "borderline-defectives" (70-79 IQ) made significantly more progress than their control group of the same low intelligence classification, while the progress of Head Start children with higher levels of intellectual functioning could not be differentiated from that of their control group with no preschool experience.

A similar finding was made by Eisenberg and Conners (1966) with a sample of Head Start children and matched controls in Baltimore. When both groups were divided into quartiles on the basis of September Peabody Picture Vocabulary test scores, the greatest improvement during the kindergarten year for Head Start as well as control children was found in the lower quartiles.

Problems in Evaluating Program Content

There are numerous problems in the assessment of intermediate range effects on Head Start programs for disadvantaged children. Studies conducted so far have

not adequately dealt with differences in the character, quality or timing of the programs under study. This is not simply a matter of neglect or error in research design. The problems involved in the adequate assessment, control and description of the several possibly significant program variables are enormous. Not only do the formal program components vary but many less formal aspects of the learning situation are inconstant from one class to the other and have been determined to be significant influences. Teacher attitudes vary between teachers and possibly for the same teacher with different groups of children. Teacher behaviors differ in relation to the quality of teacher. Teacher behaviors also vary with respect to amount and kind of attention given to pupils based on the relative position of each child in the class. These examples by no means exhaust the many program variables evaluation research methodology has yet to treat.

If and when these methodological problems have been solved, there will remain problems related to differential response to patterns of program input based upon a considerable degree on variation in the readiness, potential, and environmental support for learning in the children who are being served. There we are not only confronted with the idiosyncratic response patterns of individual children, but differential individual and group response based on the composition of the class, the child's position in the class, his previous experience and developmental level and also the quality and speed of initial response. All of these factors complicate the interpretation of evaluation data. In too many instances interpretation is even more complicated by the absence of data on such factors.

Whether one approves or not Head Start is colloquially regarded as a program designed to improve the child's chance of making it in school. In view of this fact, the heavy emphasis on cognitive development is understandable. Yet this disproportionate focus constitutes a problem in evaluation. Since all of the programs under study have not stressed under study have not stressed cognitive development as an input, it is not entirely appropriate that they be assessed on the basis of the extent to which they improve cognitive function. The mechanisms by which cognitive and affective development are interrelated are certainly not completely understood. However, it is clear that reciprocal interactions are a primary feature of these two learning modalities. To focus on one and ignore the other introduces an imbalance in assessment which may greatly distort the evaluation findings. In looking for intermediate and long term effects of intervention, it is possible that the earliest gains my be attitudinal, motivational and social. Cognitive gains may follow as a result of the affective changes more so than as a result of any specific program input. On the other hand, a specific training procedure may result in the development of a skill which becomes the vehicle upon which major strides in affective development may occur. Instrumentation for the assessment of affective development in young children is currently inadequate to permit optimal investigation of these issues in evaluation research.

Even if we were justified in the disproportionate emphasis on cognitive development in Head Start evaluation, there are problems in the assessment of intelligence which further complicate the task. The literature abounds with critiques and defenses of current psychometric instrumentation and their application for disad-

vantaged populations. The weaknesses usually discussed--cultural bias, unfamiliar measurement situations, verbal skills-- laden items, etc.--are handicaps uniformly imposed on all children from disadvantaged backgrounds when they are tested. However, the suggestion that the test instruments may be insufficiently sensitive to behavioral status and change at multiple levels of function presents a new and special problem for program evaluations. If we fail to show progress because a) the youngster has hit a test ceiling in a particular area or level of function, or b) there are important aspects of function which the instrument does not measure, or c) there is insufficient variance in the quality of functions tapped by the item, our effort at evaluation is considerably frustrated. The differences between some clinical estimates and formal measures of pupil progress suggests that psychometric instrumentation is retarding to a certain degree the sophistication of evaluation research.

While none of the studies referred to either individually or combined present an affirmation of Head Start, they do show something of the complexities involved in any follow-up evaluation. Should future follow-up studies show that present intelligence and achievement tests are, in fact, primarily geared towards measuring changes at lower levels of function, new tests for this age group and for these levels may have to be developed. It should be noted that even now various tests differ significantly in the degree of retardation which they show lower-class children to have in comparison to middle-class children; similarly, different tests show varying amounts of progress as a result of the preschool experience. It may also be found,

however, that both Head Start and kindergarten classes are often pitched at too low a level to stimulate optimal development in young children with middle or high initial levels of functioning; this too will be an important discovery. Until more carefully controlled research is directed at the many facets which may influence cognitive and social performance in young children, it will be difficult to assess the effectiveness of any Head Start or other preschool program. Fortunately this work is being undertaken with increasing quantity and quality. It may well be that within the next year or two our competence and our understanding will be more adequate than it is today.

Bibliography

Chesteen, Hillard E. (1966) "Effectiveness of the Head Start Program in Enhancing School Readiness of Culturally Deprived Children," Baton Rouge: Community Advancement Inc., Unpublished OEO report, June.

Coleman, James S. and others (1966) Equality of Educational Opportunity, National Center for Educational Statistics, Washington: Government Printing Office.

Eisenberg, Leon & Conners, C. Keith (1966) "The Effect of Head Start on Developmental Processes," Baltimore: Johns Hopkins Hospital, Unpublished OEO report.

Gordon, Edmund W. & Wilkerson, Doxey (1966) Compensatory Education for the Disadvantaged, New York: College Entrance Examination Board.

Grey, Susan & others (1966) Before First Grade: The Early Training Project for Culturally Disadvantaged Children, New York: Teachers College Press.

Kohlberg, Lawrence (1967) Assessment of a Montessori Program Delivered at American Educational Research Association, New York City.

Sigel, Irving E. & McBane, Bonnie (1966) "The Relationship between Cognitive Competence and Level of Symbolization Among Five-Year-Old Children," Detroit: Merrill-Palmer Institute. Unpublished OEO report.

U.S. Commission on Civil Rights (1967) Racial Isolation in the Public Schools, Washington: Government Printing Office.

Weikart, David P. (1964) "Perry Preschool Project Progress Report - Review of 1960-63, Ypsilanti Public Schools.

Wolff, Max & Stein, Annie (1966) "Study I: Six Months Later: A Comparison of Children Who Had Head Start, Summer 1965, with their Classmates in Kindergarten," (A Case Study of kindergartens in four public elementary schools, New York City) New York: Yeshiva University. Unpublished OEO report.

Zigler, Edward (1966) "Mental retardation: Current Issues and Approaches," in Hoffman & Hoffman (eds) Child Development Research, Russell Sage Foundation.